



[12] 发明专利申请公开说明书

[21] 申请号 02134670.4

C06F 9/06 C06F 15/16

[43] 公开日 2004 年 3 月 10 日

[11] 公开号 CN 1480859A

[22] 申请日 2002.9.3 [21] 申请号 02134670.4

[71] 申请人 鸿富锦精密工业（深圳）有限公司

地址 518109 广东省深圳市宝安区龙华镇油
松第十工业区东环二路 2 号

共同申请人 鸿海精密工业股份有限公司

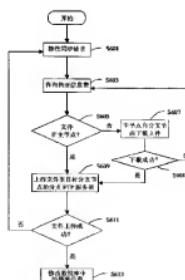
[72] 发明人 李忠一 林海洪 樊晓迪

权利要求书 3 页 说明书 6 页 附图 5 页

[54] 发明名称 分布式文件同步系统及方法

[57] 摘要

一种分布式文件同步系统及方法，该系统包括一主节点、若干分支节点以及将主节点和上述分支节点相连的网络。其中主节点是同步操作的控制中心，该主节点包括一主数据库服务器，一主 FTP 服务器及一应用程序服务器。应用程序服务器包括一同步控制模块，用于控制主节点的主 FTP 服务器与分支节点的分支 FTP 服务器间的数据传输，并更新主节点的数据库服务器中文件摘要信息。该系统的分支节点至少包括一分支数据库服务器及一分支 FTP 服务器。本发明通过主节点的应用程序服务器运行同步程序，使文件同步操作与分支节点分离，从而更易于分支节点的扩充。



1.一种分布式文件同步系统，该系统包括一主节点、若干分支节点以及将主节点和上述分支节点相连的网络，其特征在于：

分支节点包括一分支数据库服务器及一分支FTP服务器，其中：

分支数据库服务器用于存储文件摘要信息，其与主节点的主数据库服务器中的文件摘要信息保持同步；

分支FTP服务器用于存储文件，并可与主节点的主FTP服务器进行文件的传输；

主节点是同步操作控制中心，该主节点包括一主数据库服务器、一主FTP服务器及一应用程序服务器，其中：

主数据库服务器，用于存储文件摘要信息，并控制上述摘要信息与各分支数据库服务器同步；

主FTP服务器，用于存储文件，并可上传文件至分支FTP服务器及自分支FTP服务器下载文件；

应用程序服务器，用于控制同步操作进行，该服务器包括一同步控制模块，用于控制主FTP服务器自分支FTP服务器下载文件及上传文件至分支FTP服务器，并更新主数据库服务器中的文件的摘要信息。

2.如权利要求1所述的分布式文件同步系统，其特征在于，其中主节点及分支节点皆包括若干客户端，该客户端是文件使用者。

3.如权利要求1所述的分布式文件同步系统，其特征在于，其中主数据库服务器及分支数据库服务器包括一文件摘要信息表，该摘要信息表包括文件编号、存储位置、更新时间。

4.如权利要求1所述的分布式文件同步系统，其特征在于，其中主FTP服务器及分支FTP服务器中存储的文件包括文档、声音、图像及影像文件。

5.如权利要求1所述的分布式文件同步系统，其特征在于，其中主数据库服务器与分支数据库服务器间的摘要信息同步为定时进行。

6.如权利要求1所述的分布式文件同步系统，其特征在于，其中主节点应用程序服务器包括一数据库连接模块，用于连接主数据库服务器，对主数据库服务器中的数据进行查询、更新及删除操作。

7.如权利要求1所述的分布式文件同步系统，其特征在于，其中主节点应用程序服务器包括一文件传输控制模块，用于控制主FTP服务器与分支FTP服务器间的数据传输操作。

8.一种分布式文件同步方法，其特征是该方法包括以下步骤：

(a) 将若干分支节点的分支数据库服务器中文件摘要信息与主节点的主数据库服务器中的文件摘要信息同步；

(b) 第一分支节点发送文件同步请求讯息至主节点；

(c) 主节点接收该同步请求讯息，判断所请求的文件的具体位置，并获取该文件；

(d) 主节点将该文件上传至第一分支节点；

(e) 主节点在主数据库服务器中更新文件摘要信息。

9.如权利要求8所述的分布式文件同步方法，其特征在于，其中文件摘要信息包括文件编号、存储位置、更新时间。

10.如权利要求8所述的分布式文件同步方法，其特征在于，其中步骤

(a) 中的同步为定时进行。

11.如权利要求8所述的分布式文件同步方法，其特征在于，其中步骤(b)包括以下步骤：

(b1) 第一分支节点的客户端在第一分支节点的分支数据库服务器中搜索待操作的文件摘要信息；

(b2) 若文件最新版本在该第一分支节点的分支FTP服务器中，则打开该文件，若文件最新版本不在第一分支节点的分支FTP服务器上，则发送一文件同步请求讯息至主节点的同步控制模块。

12.如权利要求8所述的分布式文件同步方法，其特征在于，其中步骤

(c) 包括以下步骤：

(c1) 主节点在其主数据库服务器中搜索文件摘要信息，获得文件最新版本位置；

(c2) 若文件最新版本在主节点的主FTP服务器中，则执行步骤(d)；

(c3) 若文件最新版本在第二分支节点的分支FTP服务器中，则主节点的主FTP服务器首先自该第二分支节点下载文件至主节点的主FTP服务器。

-
- 13.一种分布式文件同步方法，其特征是该方法包括以下步骤：
- (a) 将若干分支节点的分支数据库服务器中的文件摘要信息与主节点的主数据库服务器中的文件摘要信息同步；
 - (b) 第一分支节点的客户端在该第一分支节点的分支数据库服务器中搜索一待操作文件摘要信息；
 - (i) 若待操作文件最新版本在该第一分支节点的分支FTP服务器中，则打开该待操作文件，操作结束；
 - (ii) 若待操作文件最新版本不在第一分支节点的分支FTP服务器上，则发送一文件同步请求讯息至主节点的同步控制模块；
 - (c) 同步控制模块搜索主节点的主数据库服务器中的摘要信息，获得待操作文件最新版本位置；
 - (i) 若待操作文件最新版本在主节点的主FTP服务器，则主节点将文件上传至第一分支节点；
 - (ii) 若待操作文件最新版本在一第二分支节点的分支FTP服务器中，则主节点首先自第二分支节点下载待操作文件至主节点，然后将该待操作文件上传至第一分支节点；
 - (d) 主节点在主数据库服务器中更新文件摘要信息。

分布式文件同步系统及方法

【技术领域】

本发明涉及一种文件同步系统及方法，特别涉及一种分布式文件同步系统及方法。

【背景技术】

随着经济全球化的发展趋势，出现了一大批跨地区、跨国界的大型企业集团。这些企业集团的分支机构分布于世界各地，而各地分支机构需要共享信息以联合工作。因此，如何对其数据进行管控，以实现企业信息资源共享，已成为该类企业营运的一个重要组成部分。而信息技术的发展，使得跨机构、跨地域的数据访问与协作成为可能。目前，为了提高数据访问的效率以及减小故障的损失，企业一般采用分布式数据管理系统。

分布式数据管理系统指物理空间分散、逻辑上通过网络连接在一起的数据存储、管理系统，在该系统中，用户能像访问本地存储区的数据一样使用远程存储区中的数据。如2002年6月12日公开的公开号为CN1353836的名为“分布式环境中的文件管理方法和系统”的中国专利，该专利在客户端设置文件控制器，管理本地文件及控制网络信息交换，并生成包含文件存储位置的元索引以标志网络信息。在用户在某客户端进行文件访问时，通过元索引获得文件存储位置，并将文件复制到本地终端，同时修改元索引信息。该专利解决了特定终端负载集中及文件更新同步的问题，但该系统却存在着客户端数据重复存储、同步过程通讯量较大、系统不易扩充等问题。

【发明内容】

本发明提供一种分布式文件同步系统及方法，通过主节点的应用程序服务器运行同步程序，使文件同步操作与分支节点分离，以解决系统不易扩充的问题。

该系统包括一主节点、若干分支节点以及将主节点和上述分支节点相连的网络。分支节点包括一分支数据库服务器及一分支FTP服务器，其中

分支数据库服务器用于存储文件摘要信息，其与主节点的主数据库服务器中的文件摘要信息保持同步；分支FTP服务器用于存储文件，并可与主节点的主FTP服务器进行文件的传输。主节点是同步操作控制中心，该主节点包括一主数据库服务器、一主FTP服务器及一应用程序服务器，其中主数据库服务器，用于存储文件摘要信息，并控制上述摘要信息与各分支数据库服务器同步；主FTP服务器，用于存储文件，并可上传文件至分支FTP服务器及自分支FTP服务器下载文件；应用程序服务器，用于控制同步操作进行，该服务器包括一同步控制模块，用于控制主FTP服务器自分支FTP服务器下载文件及上传文件至分支FTP服务器，并更新主数据库服务器中的文件的摘要信息。

本发明还提供一分布式文件同步方法。该方法包含以下步骤：(a) 将若干分支节点的分支数据库服务器中文件摘要信息与主节点的主数据库服务器中的文件摘要信息同步；(b) 第一分支节点发送文件同步请求讯息至主节点；(c) 主节点接收该同步请求讯息，判断所请求的文件的具体位置，并获取该文件；(d) 主节点将该文件上传至第一分支节点；(e) 主节点在主数据库服务器中更新文件摘要信息。

通过本发明，可使存储于不同地域的文件在相对较小通讯量的情形下能够准确同步，并且使得系统的分支节点易于扩充。

【附图说明】

图1是本发明分布式文件同步系统及方法的实施环境图。

图2是本发明分布式文件同步系统及方法的分支节点结构图。

图3是本发明分布式文件同步系统及方法的主节点结构图。

图4是本发明分布式文件同步系统及方法的数据库服务器中文件摘要信息表。

图5是本发明分布式文件同步系统及方法的客户端读取文件流程图。

图6是本发明分布式文件同步系统及方法的文件复制流程图。

【具体实施方式】

如图1所示，是本发明分布式文件同步系统及方法的实施环境图。该系统由分散在不同地理位置的企业信息系统节点组成，其具体包括一主节点110及若干个分支节点120、130（图中只画出二个）。这些节点通过企业

内部网140相连，且都采用三层（或多层）结构：一数据库服务器（113、123或133），一系列的应用服务器（112、122或132），以及分散的客户端（111、121或131）。以下将以主节点110为例，介绍各节点的组成。

主节点110包括一数据库服务器113，一应用服务器系统112及若干客户端111（图中仅画出一个）。数据库服务器113用于存储企业相关的所有数据，并负责管理对数据库的读写和维护，能够迅速执行大量数据的更新和检索。应用服务器系统112包括了企业信息系统中核心的和易变的企业逻辑（规划、运作方法、管理模式等），它的功能是接收输入，处理并返回结果。客户端111为系统应用的用户接口，负责接收用户输入及向用户展示系统处理结果。分支节点120、130的组成结构与主节点110相同。

图2是本发明分布式文件同步系统及方法的分支节点120的结构图。分支节点120的客户端121包括一客户端代理程式201，用于将系统操作封装，使系统操作与用户操作分离，简化了用户操作。该客户端代理程式201包括一用户操作模块202、一数据库操作模块203及一文件操作模块205。分支节点120的应用服务器系统122包括一FTP（File Transfer Protocol）服务器207及一应用程序服务器209。

用户操作模块202为用户提供一图形用户接口（Graphical User Interface, GUI），方便用户输入。该用户操作模块202与数据库操作模块203及文件操作模块205相连，将系统对FTP服务器207及数据库服务器123的操作结果展示给用户。文件操作模块205与FTP服务器207连接，用于向FTP服务器207发送对文件的操作命令，该命令包括文件打开、关闭、编辑、上传、下载、删除、变更及增加或删除目录夹等操作。文件操作模块205还与数据库操作模块307相连，用于处理数据库操作模块307传送的数据。数据库操作模块203与应用程序服务器209相连，用于生成对数据库服务器123中的数据进行查询、更新及删除等操作命令。数据库操作模块203还负责将数据库服务器123的操作结果传送至文件操作模块205。

FTP服务器207用于存储文件，如文档、图像、声音及影像文件等。该FTP服务器207与企业内部网140相连，可通过企业内部网140与其它节点的FTP服务器进行文件传输操作。应用程序服务器209用于连接客户端121与数据库服务器123，将客户端121的数据库操作模块203对数据库操作的命

令传递至数据库服务器123执行，并将数据库服务器123执行结果传送给数据库操作模块203。

数据库服务器123存储了企业文件（如声音、图像、文档及影像等文件）的摘要信息，该摘要信息包括文件编号、存储位置及目录、更新日期等信息。该数据库服务器123中的文件摘要信息定时与主节点110的数据库服务器113中的文件摘要信息同步。

图3是本发明分布式文件同步系统及方法的主节点110的结构图。主节点110的客户端111及数据库服务器113的功能、组成与结构与分支节点120相同（见图2）。

主节点110的应用服务器系统112包括一FTP服务器309及一应用程序服务器310。FTP服务器309用于存储文件，包括文档、图像、声音及影像等。该FTP服务器309与企业内部网140相连，可通过企业内部网140与其它节点的FTP服务器进行文件传输操作。应用程序服务器310是该文件同步系统的控制中心，其包括一同步控制模块313、一文件传输控制模块311及一数据库连接模块315。同步控制模块313用于控制同步操作的执行，其与文件传输控制模块311及数据库连接模块315相连。该同步控制模块313可根据同步请求生成若干个同步操作命令，这些同步操作命令包括数据传输命令及更新数据库命令，然后同步控制模块313发送命令至文件传输控制模块311及数据库连接模块315执行。文件传输控制模块311可接收FTP服务器309的同步请求并将该请求发送至同步控制模块313以生成同步控制命令。同时，文件传输控制模块311还负责解读同步控制模块313的同步命令，并将该同步命令编译为FTP服务器309可执行的指令。数据库连接模块315可执行同步控制模块313的数据库更新命令，将数据同步信息如位置变更、时间更新等信息更新。

图4是本发明分布式文件同步系统及方法的数据库服务器113（或213、313）中文件摘要信息表400。该摘要信息表400包括文件编号401、文件标题403、文件位置405、文件目录407及更新时间409。其中，文件编号401是对应文件在企业信息系统中的统一编号，每一文件有且仅有一唯一的编号。文件标题403是文件的标题，其可简要描述该文件的内容。文件位置405是存储文件所在的FTP服务器的IP地址，如10.153.24.126。文件目录407

是文件在FTP服务器中的目录。更新时间409是上一次对文件操作的具体时间。分支节点210（或310）的数据库服务器213（或313）中的摘要信息表400定时与主节点110的数据库服务器113中的摘要信息表400同步。例如，可在主节点113的数据库服务器113中设定每三分钟执行一次摘要信息表400同步步骤。

图5是本发明分布式文件同步系统及方法的客户端121读取文件流程图。首先，用户通过用户操作模块202的GUI向系统发送对一文件的操作请求（步骤S501），该操作可以为阅读或编辑该文件，该文件对应唯一编号。数据库操作模块203将该请求编译为数据库查询语句，并经由应用程序服务器209传递至数据库服务器123。数据库服务器123执行该查询语句，在文件摘要信息表400的文件编号401栏中查询符合条件的表项（步骤S503）。然后，数据库服务器113将编号符合条件的表列发送至文件操作模块205。

文件操作模块205读取表列讯息，并搜索摘要信息表400表列的文件位置405栏（步骤S505）。若在该栏中无本节点200的FTP服务器207的IP地址，则表示该文件不在该节点，则文件操作模块205经由FTP服务器207向主节点110同步控制模块313发出数据复制请求，该数据复制请求仅含数据编号（步骤S507），该步骤将在图6详细描述。若本地有该数据，即文件位置405栏有本节点200的FTP服务器207的IP地址，则文件操作模块205进一步比较摘要信息表400表列的更新日期栏（步骤S501）。若对应FTP服务器207的IP地址的表列的更新日期早于其余任一表列的更新日期，即该数据最新版本不在本节点200，则执行步骤S507。若数据最新版本在本节点200的FTP服务器207，则文件操作模块205将该文件直接显示给用户。

图6是本发明分布式文件同步系统及方法的文件复制流程图。下文例举分支节点120与主节点110的同步进行说明。首先，主节点110的FTP服务器309接收分支节点120的FTP服务器207发送的文件同步请求（步骤S601）。该文件同步请求经由文件传输控制模块311传递至同步控制模块313。同步控制模块313通过数据库连接模块315在数据库服务器113的摘要信息表400中查询数据复制请求中数据编号，并将查询结果表列传递回同步控制模块313（步骤S603）。

同步控制模块313根据摘要信息表400中的文件位置405栏及更新日期409栏判断该文件最新版本位置（步骤S605）。若数据最新版本不在主节点110的主FTP服务器309中，则同步控制模块313发送数据下载命令至文件传输控制模块311，该数据下载命令包括数据编号、位置及目录，文件传输控制模块311将该下载命令编译成FTP服务器309可直接执行的下载指令，并发送至FTP服务器309，FTP服务器309执行下载指令，从文件最新版本所在节点（如分支节点130）的FTP服务器下载文件最新版本（步骤S607），同时生成一响应讯息，通知同步控制模块313文件已下载（步骤S608），若文件下载失败，则返回步骤S603。若文件最新版本在主节点110的FTP服务器309上，则同步控制模块313生成一文件上传命令，并将该命令发送至文件传输控制模块311，经文件控制模块311编译数据上传指令被FTP服务器309执行，相应数据被传递至原请求节点120的FTP服务器207（步骤S609）。若文件上传成功，FTP服务器309将一上传成功讯息经由文件传输控制模块311返回同步控制模块313（步骤S311），同步控制模块313将修改数据库服务器113中的摘要讯息表400。若文件上传失败，则返回步骤S601。

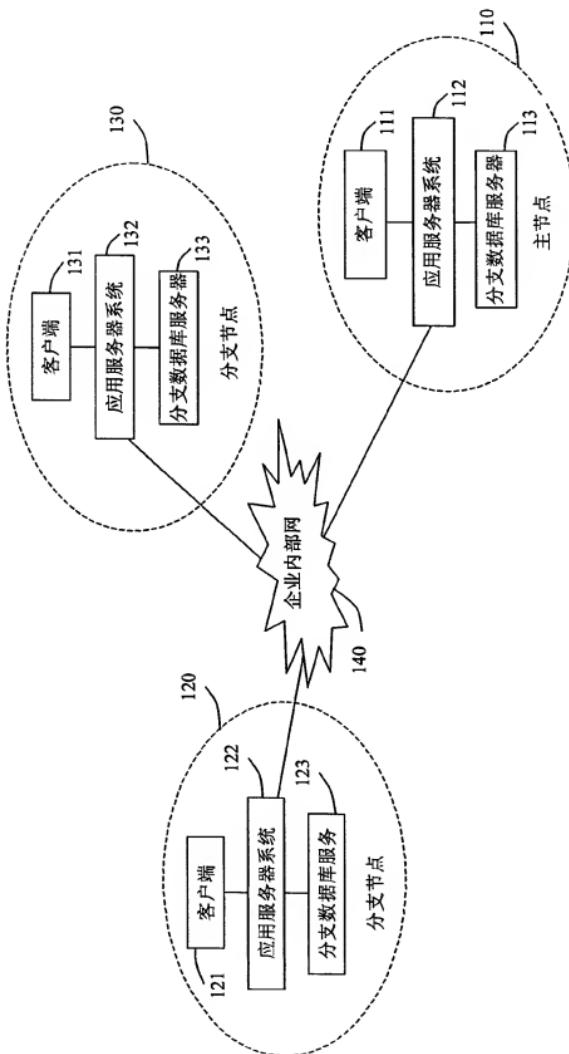


图 1

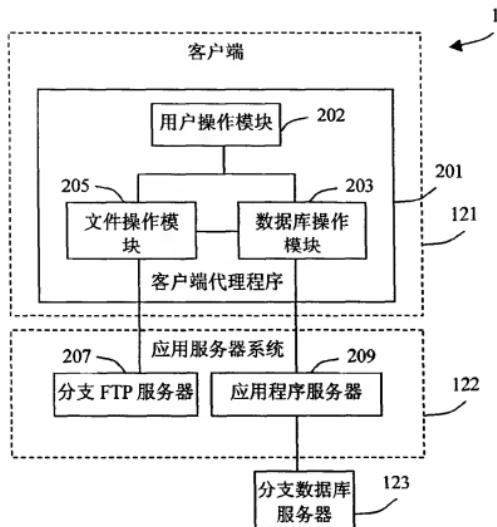


图 2

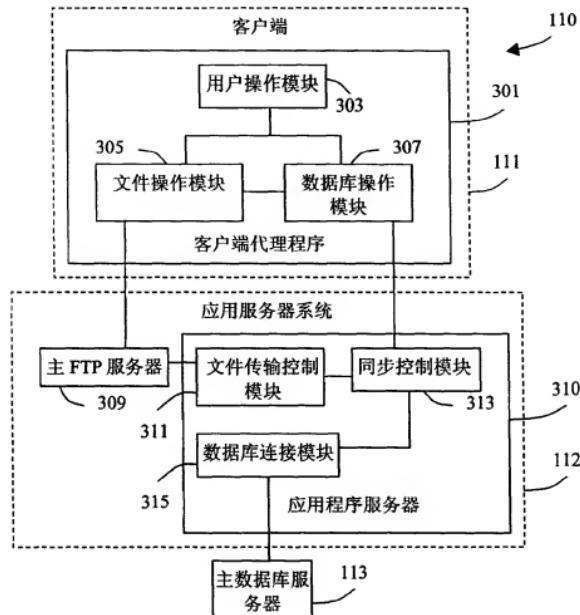


图 3

A table representing file metadata:

文件编号 (File Number)	文件标题 (File Title)	文件位置 (File Location)	文件目录 (File Directory)	更新时间 (Update Time)	...

Column numbers are indicated by arrows:

- 401: Points to the first column (File Number)
- 403: Points to the second column (File Title)
- 405: Points to the third column (File Location)
- 407: Points to the fourth column (File Directory)
- 409: Points to the fifth column (Update Time)
- 400: Points to the ellipsis column (...)

图 4

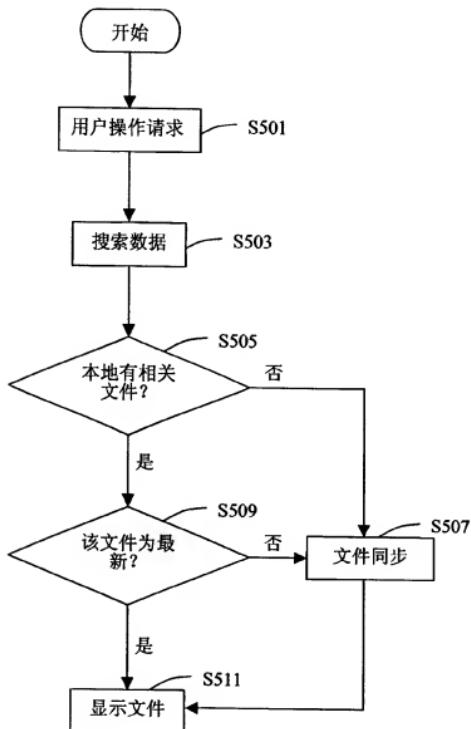


图 5

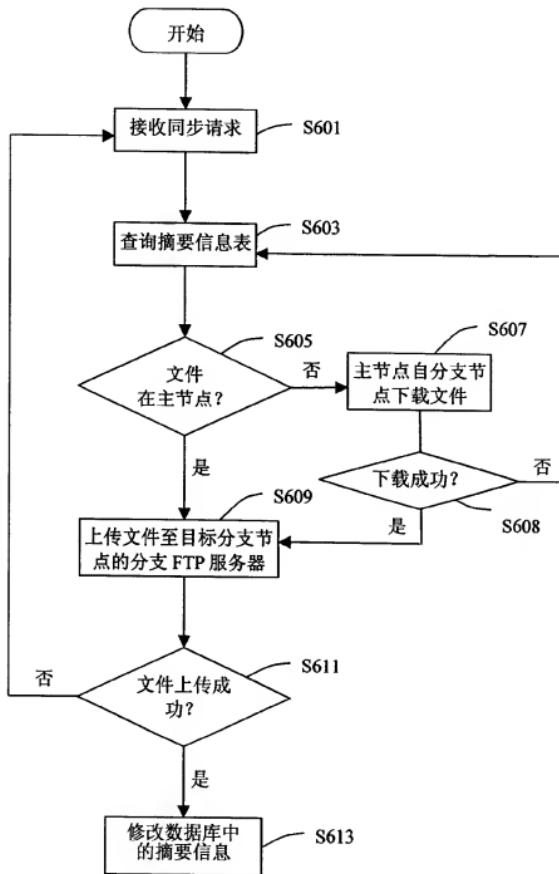


图 6